DOCUMENT RESUME

ED 172 980 RC 011 353

AUTHOR Martin, Wayne H.
TITLE Student Achievement in Rural Schools: A View from the National Assessment Data.
SPONS AGENCY Bureau of Elementary and Secondary Education (DHEW/OE), Washington, D.C.; Department of Agriculture, Washington, D.C.; National Inst. of Education (DHEW), Washington, D.C.; Office of the Assistant Secretary for Education (DHEW), Washington, D.C.
PUB DATE May 79
NOTE 32p.; Paper presented at the Rural Education Seminar (College Park, Maryland, 29-31 May 1979); Some pages may not reproduce due to small print size

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Achievement; Comparative Analysis; Demography; *Educational Assessment; Elementary Secondary Education; Family Environment; Grade 4; Grade 8; Grade 11; *National Surveys; Rural Areas; Rural Education; *Rural Schools; *Rural Youth
IDENTIFIERS *National Assessment of Educational Progress

ABSTRACT

The National Assessment of Educational Progress (NAEP) was designed to measure knowledge, skills, and attitudes of young Americans at various ages in 10 learning areas, and to measure educational attainment over time. Community categories used in NAEP research were High and Low Metro, Urban Fringe, Main Big City, Medium City, Small Places (population under 25,000), and Extreme Rural Areas (population under 10,000). Rural students at ages 9, 13, and 17 were enrolled at near the national average grade level. The home environment of 9 year old rural students was similar to that of Urban Fringe students. At ages 13 and 17, the home environments were similar to Main Big City home environments. The baseline data trend, substantiated by change data, was toward improved rural performance, to the point of reaching national performance levels for some ages in science, reading, functional literacy, and social studies. Exercise-by-exercise data should be examined to isolate strengths and weaknesses of rural students in various learning areas. The federal government should explore the possibility of conducting a migrant children assessment. NAEP should be provided with the necessary resources to increase the sample size for rural students to provide more detailed information about their educational achievement. (Author/SB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Wayne H. Martin

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED172980

STUDENT ACHIEVEMENT IN RURAL SCHOOLS: A VIEW FROM THE NATIONAL ASSESSMENT DATA

Prepared for the
Rural Education Seminar

May 29-31, 1979

Paper presented at the Rural Education Seminar, College Park, Maryland, 29-31 May 1979. Seminar sponsored by the U. S. Department of Health, Education and Welfare (Office of the Assistant Secretary of Education, Office of Education's Bureau of Elementary and Secondary Education, and the National Institute of Education's Program on Educational Policy and Organization) and the U. S. Department of Agriculture (Science and Education Administration).

By: Wayne H. Martin
Education Commission of the States, Denver, CO

CE 011353

INTRODUCTION

The purpose of this paper is to examine the National Assessment of Educational Progress (NAEP) data with regard to the performance of rural students. For the reader who is familiar with National Assessment, it is not necessary to read the first section of the paper ("Overview of National Assessment"). The second section of the paper presents National Assessment's definition of rural students and the other size-and-type of community categories that NAEP uses to report student performance. It also provides some demographic and soicological data about the composition of the various size-and-type of community categories that the reader will find useful in considering the remainder of the paper. The third section of the paper examines baseline data from various assessment years for rural students in relation to the nation and other size-and-type of community categories. The fourth section of the paper considers the performance of rural students in terms of National Assessment's change over time data. The final section of the paper provides a brief summary of the performance of rural students and recommendations for further research that is needed in the area of student achievement in rural areas.

1. OVERVIEW OF NATIONAL ASSESSMENT

Purpose

The National Assessment of Educational Progress was established to determine what young Americans know and can do. Four age levels were selected for assessment: age 9, when most students have completed their primary education; age 13, when most students have completed their intermediate education; age 17, when most students have completed their secondary education; and ages 26 to 35, when most individuals have completed their formal post-secondary education. Specifically, the assessment program was designed to measure the knowledge, skills, and attitudes possessed by young Americans at these key points in the educational system and to monitor changes (growth or decline) in their educational attainments over time. Ten learning areas were selected for assessment purposes: art, career and occupational development, citizenship, literature, mathematics, music, reading, science, social studies, and writing. The three in-school ages are assessed annually in one or more learning areas. Although the assessment of young adults was initially planned to be on an annual basis, it has been necessary to survey them on a periodic basis because of budgetary constraints.

Methodology

To measure the nation's educational progress, National Assessment estimates the percentage of respondents at each age level who are able to answer a question acceptable or perform a task. Each question or task (called an exercise) reflects a previously defined educational goal or objective. The objectives for each learning area are developed (or redeveloped after the first assessment of the area) using a

consensus approach which involves scholars in the field, educators, and concerned lay people. The exercises are administered to scientifically selected samples which take into account size of community and socioeconomic status. The samples are designed so that sound inferences can be made about the populations from which the samples were selected.

National Assessment does not develop or use scores for individual respondents. Rather, it determines how each age level performs on specific exercises and, within each age level, how groups of individuals (based on demographic and sociological variables) perform. Thus, it is not necessary for each respondent to take every exercise. The exercises are divided into booklets or packages, and each in-school respondent takes only one package. Since the samples for the different packages are statistically equivalent, group comparisons can be made across packages. This allows NAEP to assess performance on far more exercises in a learning area than would be possible in the usual 50 minute situation; it also provides broader coverage of the objectives for each learning area. The reporting variables are presented in Table 1.

Although the assessment utilizes multiple-choice exercises (and these predominate in some of the learning areas), many open-ended or free response exercises are also utilized. These range from exercises requiring a few words as an answer to those requiring a long essay. The assessment packages are considered to be "consumable" booklets and respondents are required to do all work within them. This has allowed for the development of detailed scoring analyses for many of the

open-ended exercises. During development, exercise writers attempt to develop exercise formats that provide the best and most direct measure of the objective being assessed. For example, the writing assessment requires respondents to demonstrate their writing skills; the art assessment requires students to demonstrate their drawing and design skills; and the music assessment has required respondents to demonstrate their musical abilities. To determine the effect of hand-held calculators on students' abilities to compute, the recent mathematics assessment provided calculators for some respondents to use during the assessment. Individual interviews, the manipulation of apparatus to solve a problem, and observations of the respondents' problem-solving techniques are also used to supplement the usual paper and pencil measures.

TABLE 1. National Assessment Reporting Variables

Variable	Categories	
Age	9-year-olds 13-year-olds 17-year-olds Adults (26-35 years)	
Region	Northeast	(Delaware, Connecticut, Maine, New Hampshire, Rhode Island, Vermont, District of Columbia, Maryland, Massachusetts, New Jersey, Pennsylvania, New York)
	Southeast	(Arkansas, Florida, Virginia, West Virginia, Alabama, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee)
	Central	(Iowa, Kansas, Nebraska, North Dakota, South Dakota, Minnesota, Missouri, Illinois, Indiana, Michigan, Ohio, Wisconsin)
	West	(Alaska, Hawaii, Idaho, Montana, Nevada, Wyoming, Utah, Arizona, Oregon, Colorado, New Mexico, Oklahoma, Texas, California, Washington)
Sex	Male, Female	
Race	Black, White, Other	
Size and Type of Community	High Metropolitan, Low Metropolitan, Extreme Rural, Main Big City, Urban Fringe, Medium City, Small Places	
Level of Parental Education	No High School, Some High School, Graduated High School, Post High School	

National Assessment considers positive attitudes toward or opinions about the various learning areas to be important educational attainments. Therefore, affective exercises and attitude survey questions are included in the assessment of most learning areas.

Because individuals are not ranked according to their performance on the assessment materials, National Assessment does not emphasize the use of exercises with high discrimination power. The aim of the assessment is to describe educational attainments. This is best accomplished if the exercises included in an assessment cover a broad range of difficulty, from very easy tasks to the most difficult.

A representative national sample of public and private schools is assessed each year. Participation in the assessment is on a voluntary basis for the schools selected for each year's assessment sample. Over 90 percent of the selected schools agree to participate in the assessment. The data collection is geared to the school year. Thirteen-year-olds are assessed during October through mid-December; nine-year-olds during January and February; and seventeen-year-olds are assessed from March through mid-May.

In order to guarantee uniformity of data collection, the assessment is administered by specially trained personnel who are stationed all over the country. These individuals do a portion of the assessment administrations themselves and hire and train local persons for the remainder. Thus a school district is asked only to provide the students selected in the sample and space for the assessment administration. A student may opt not to participate if he or she wishes. Most assessment materials

are administered to groups of 16 to 20 students using a paced tape technique. The administrator plays a recorded set of directions and then each exercise (including the answer choices for multiple-choice exercises) is read aloud on the tape. The tape provides silence for a period of time sufficient to answer each question. This technique insures uniform presentation of directions and provides respondents with a reading problem an opportunity to hear the exercise as they are reading it. The technique is not used for the reading assessment or for individually administered materials. A maximum of 50 minutes is required from an individual student participating in the assessment. These techniques were designed to minimize the disruption to a participating school.

Student names are not recorded on the assessment booklets. They are used only for the selection of the local sample. Thus, students are guaranteed anonymity. In addition, no reports are developed for school districts or for states. The sample design does not permit making inferences to any geographic unit smaller than a region.

National Assessment reports results on all exercises and, in addition, publishes the entire text of approximately half of the exercises. This allows for a detailed understanding of the assessment's purposes and construction; it also allows for greater flexibility in interpretation or further analysis of the data. The unpublished items are kept secure and used in subsequent assessments of the learning area to measure changes in performance over time.

During the ten years that it has been collecting data, National Assessment has conducted and reported the initial assessment for 9-, 13-, and 17-year-olds in each of the ten learning areas selected for assessment. It has also developed, conducted, and reported the first measurement of change in five learning areas and the second measurement of change in one learning area (refer to Table 2). By the end of the current contract, National Assessment will have reported the first measurement of change in mathematics, conducted the first measurement of change for art and music and the second measurement of change for writing, and initiated data collection for the second measurement of change in reading and the first measurement of change in literature.

TABLE 2

National Assessment Timetable, 1969-79
(expressed in school years)

Learning Area	Initial Assessment	First Measurement of Change	Second Measurement of Change
Art	1974-75	1978-79	
Career and Occupational Development	1973-74		
Citizenship	1969-70	1975-76	
Literature	1970-71	1979-80	
Mathematics	1972-73	1977-78	
Music	1971-72	1978-79	
Reading	1970-71	1974-75	1979-80
Science	1969-70	1972-73	1976-77
Social Studies	1971-72	1975-76	
Writing	1969-70	1973-74	1978-79

National Assessment also recognizes that special needs for data not included in the ongoing assessment frequently arise. To address such needs, the program has added special purpose "probes". A probe is designed to be a small scale assessment of a specific topic and is usually administered to only one age group. It is designed to provide a "snapshot" of the topic and is not necessarily designed to measure change.

National Assessment has administered probes in the areas of basic life skills (1977) and consumerism (1978) for 17-year-olds and in the areas of health and energy (1977) for adults. In addition, the assessment program administered, analyzed, and reported the assessment of functional literacy of 17-year-old students under contract to the Right-to-Read Program. Probes provide National Assessment with the flexibility to address timely educational topics without disrupting its ongoing task of monitoring the nation's educational progress.

2. NATIONAL ASSESSMENT'S DEFINITION OF RURAL STUDENTS

As was mentioned earlier, National Assessment reports the performance for the nation at each age level and within an age level, for groups of individuals based on demographic and sociological variables. One of the reporting variables for NAEP is size-and-type of community (STOC). The categories within this variable are defined by the size of community in which a student's school is located and by an occupational profile of the area the school serves, as judged by the school's principal. All population sizes are based on data from the 1970 Census. The following categories were used in the initial assessment of all ten learning areas and the second assessments of science, reading, and writing.

Extreme rural. These schools are in areas where a high proportion of the residents are farmers or farm workers. At least some of the enrollment is from open country or places less than 2,500 population; no enrollment is from places greater than 10,000, and none is from suburbs of large cities.

Low Metro (low-socioeconomic or impoverished urban). These schools serve neighborhoods where a high proportion of the residents are on welfare or are not regularly employed. They are located in cities or residential areas of big cities with populations greater than 200,000.

High Metro (high-socioeconomic or affluent urban and suburban communities). A high proportion of the residents in these areas are professional or managerial. The schools are in big cities or residential areas of cities with populations greater than 200,000.

Main Big City. Schools in this category are located in big cities of population greater than 200,000 but not included in either the low-metro or high-metro categories.

Urban Fringe. These schools are in the urbanized areas near big cities of size greater than 200,000 and not included in either the low-metro or high-metro categories.

Medium City. These schools are in cities with populations between 25,000 and 200,000 that are not urbanized areas near big cities.

Small Places. Schools in this category are located in open country or are from places with populations of less than 25,000, not including those in the extreme-rural category.

For the reader who is either familiar with or the proud possessor of some of the initial NAEP reports, it should be noted that some of the STOC category names have changed. However, the definitions have remained the same.

Old STOC	Current STOC
Rural	Extreme Rural
Inner City	Low Metro
Affluent Suburb	High Metro
Rest of Big City	Main Big City
Suburban Fringe	Urban Fringe
Medium City	Medium City
Small City	Small Places

These changes were made because the previous category names were both technically misleading and politically value-laden.

In an attempt to provide more useful information for secondary analyses of National Assessment data and for secondary users of the released exercises, NAEP has reported the results from the second assessments of citizenship and social studies and the third assessment of science in terms of a Size of Community variable and a Type of Community variable. It should be noted that the categories within the two variables are not mutually exccusive. For example, students included in the extreme rural category for the type of community variable would be placed in the smaller places cagegory with other students for the size of community variable. The definitions used for the three assessments for these two variables are as follows:

Size of Community

Big City. Students in this group attend schools within the city limits of cities having a population over 200,000.

Fringes around big cities. Students in this group attend schools within metropolitan areas (1970 U.S. Bureau of the Census urbanized areas) served by cities having a population greater than 200,000 but outside the city limits.

Medium city. Students in this group attend schools in cities having a population between 25,000 and 200,000 and not classified in the fringes-around-big-cities category.

Smaller places. Students in this group attend schools in communities having a population less than 25,000 and not classified in the fringes-around-big-cities category.

Type of Community

These communities are defined by an occupational profile of the area served by a school as well as by the size of the community in which the school is located.

Advantaged-urban communities (high metropolitan). Students in this group attend schools in or around cities with a population greater than 200,000 where a high proportion of the residents are in professional or managerial positions.

Disadvantaged-urban communities (low metropolitan). Students in this group attend schools in or around cities with a population greater than 200,000 where a high proportion of the residents are on welfare or are not regularly employed.

Extreme rural. Students in this group attend schools in areas with a population under 10,000 where most of the residents are farmers or farm workers.

It should be noted that despite this change in reporting, the definition of the extreme rural category has remained constant for all assessments to date.

Among many individuals, there is a mistaken assumption that the problems of the low metro areas are very similar to those found in rural areas as the populations served by both are similar. Perhaps this assumption has been inadvertantly spread by the media. However, when one examines the cross-tabulation of race by STOC categories, presented in Table 3, one sees that the extreme rural category (as defined by National Assessment) is proportionately most similar to either the urban fringe or high metro STOC categories. While low metro and extreme rural students may have similar problems, the similarities would not be the result of their being the same basic population.

TABLE 3. Sample Percentages of White, Black, Hispanic and Other By Size and Type of Community

	Column Percentages	Row Percentages				
		White	Black	Hispanic	Other	Total
Age 9						
Extreme Rural	9.2	87.2	6.7	4.0	2.1	100.0
Low Metro	7.2	35.1	48.3	14.7	1.9	100.0
High Metro	10.2	87.4	8.1	2.2	2.3	100.0
Main Big City	9.5	63.2	22.8	11.1	2.9	100.0
Urban Fringe	10.6	84.9	8.0	4.8	2.3	100.0
Medium City	14.7	79.1	17.1	3.0	0.8	100.0
Small Place	38.6	88.8	7.2	2.5	1.5	100.0
National	100.0	80.4	13.5	4.7	1.4	100.0
Age 13						
Extreme Rural	10.3	85.2	5.5	6.0	3.3	100.0
Low Metro	7.2	31.3	55.0	11.9	1.8	100.0
High Metro	11.0	92.2	3.9	2.0	1.9	100.0
Main Big City	6.8	65.0	24.0	8.7	2.3	100.0
Urban Fringe	9.6	86.1	8.5	4.3	1.1	100.0
Medium City	13.4	77.5	16.5	5.0	1.0	100.0
Small Place	41.7	88.8	8.2	1.8	1.2	100.0
National	100.0	82.1	12.3	4.3	1.3	100.0
Age 17						
Extreme Rural	8.1	93.1	1.3	3.8	1.8	100.0
Low Metro	8.5	37.1	48.7	11.7	2.5	100.0
High Metro	10.2	93.3	3.5	1.9	1.3	100.0
Main Big City	6.9	76.6	16.0	5.6	1.8	100.0
Urban Fringe	13.9	93.8	3.3	1.8	1.1	100.0
Medium City	14.8	78.8	15.3	4.6	1.3	100.0
Small Place	37.6	90.5	7.1	1.5	0.9	100.0
National	100.0	83.9	11.1	3.7	1.3	100.0

Among the background data collected about respondents by National Assessment is the grade in school in which a respondent is currently enrolled. Over the years, NAEP has found that this variable does effect performance in the assessment. Traditionally, National Assessment has found that the majority of students (70 to 75%) are enrolled in 4th grade at age 9, 8th grade at age 13, and 11th grade at age 17. The 4th, 8th, and 11th grades represent modal grades for the assessment, and data have been presented in various reports for these grades. However, in preparing a report dealing with Hispanic student performance in five learning areas, NAEP found that many Hispanic students were in grades lower than generally appropriate for their age levels. In analyzing the results of the 1974-75 reading assessment, it was determined that 29 percent of the 9-year-old Hispanic students were enrolled in 3rd grade, 37 percent of the 13-year-old Hispanic students were enrolled in 7th grade, and 36 percent of the 17-year-old Hispanic students were enrolled in 10th grade. Thus, proportionally fewer Hispanic students had been exposed to curricula typically taught at the higher grades. This could account for their lower levels of performance on the assessment.

It was decided to examine the STOC categories to determine if such "surprises" were hidden in the assessment data. Table 4 presents the percent of students who are enrolled in less than the modal grade (grades 4, 8, and 11) at each age level for the nation and the STOC categories. Thus, at age 9, 23.4 percent of the 9-year-olds nationwide are enrolled in less than 4th grade. Conversely, this means that nationwide 76.6 percent of the 9-year-olds are enrolled in 4th grade or above. While none of the STOC categories have the dramatic differences found for Hispanic students, the differences among the categories are interesting. High metro students have a higher percentage

enrolled at or above the modal grade for all three ages. Among low metro students, there are approximately 4.5 percent more than the national average enrolled at or above grade level at age 9 but 3 percent less age age 13 and 12.5 percent less at age 17 enrolled at or above the modal grade than the national average. At age 9, the rural students appear to be quite close to the national average; at age 13, slightly over 4 percent more rural students are below the modal grade than the national average; and at age 17, a little more than 3 percent of the rural students are enrolled at or above the modal grade.

TABLE 4. Percent Enrolled at Less Than Modal Grade

	Age 9	Age 13	Age 17
Nation	23.4	26.5	14.2
Extreme Rural	23.4	30.[illegible]	10.9
Low Metro	18.9	29.4	26.7
High Metro	12.9	21.6	9.7
Main Big City	15.5	26.7	9.7
Urban Fringe	21.4	22.3	9.7
Medium City	22.0	27.7	14.6
Small Places	26.3	27.5	14.5

National Assessment also collects self-report information from respondents about their home environment. The four questions included in the home environment variable are: 1) Does your family get a newspaper regularly? 2) Does your family get any magazines regularly? 3) Are there more than 25 books in your home? and 4) Is there an encyclopedia in your home?. The response choices to each of the questions are "Yes",

"No", and "I don't know". It was decided to examine the response patterns to the home environment variable to determine if rural students differ dramatically from the other STOC categories. The data are presented in Table 5. It is interesting to note that the response patterns for rural students parallel fairly closely those of the urban fringe students at age 9 and those of main big city students at ages 13 and 17. While the rural students do not differ dramatically from the national response patterns, the response patterns for low metro students do differ dramatically from those of the nation and other STOC categories.

It should be noted that National Assessment does not implement any special procedures to collect data on the children of migrant farm workers. If these children are currently enrolled in a school selected for participation in the assessment, they have the same probability of being selected as an assessment respondent as any other student enrolled in that school who meets the age eligibility definition. It is not known to what extent such children are included or excluded from the assessment data.

3. BASELINE ASSESSMENT DATA

Before presenting the data for this section, it is necessary to provide some background information concerning the data. The term "baseline" was originally defined for assessment purposes as the data from the initial assessment of a learning area. The data was to serve as the "baseline" against which changes in performance over time would be measured. Yet as National Assessment began collecting change data, the definition of "baseline" began to evolve to recognize the fact that every assessment of a learning area was a "baseline". As a result, National Assessment began to think of

TABLE 5. Response Patterns to Home Environment Questions by STOC Categories

	"Yes" to less than 3 Home Environment Questions	"Yes" to 3 of the Home Environment Questions	"Yes" to all 4 of the Home Environment Questions	
Age 9				
Nation	33.4	32.3	34.2	100.0
Extreme Rural	35.5	31.8	32.7	100.0
Low Metro	59.8	25.0	15.2	100.0
High Metro	26.5	27.8	45.7	100.0
Main Big City	42.7	30.0	27.3	100.0
Urban Fringe	33.4	34.1	32.5	100.0
Medium City	25.4	34.7	39.9	100.0
Small Places	30.6	34.3	35.1	100.0
Age 13				
Nation	18.4	26.2	55.4	100.0
Extreme Rural	21.9	28.5	49.6	100.0
Low Metro	47.7	23.0	29.3	100.0
High Metro	7.2	19.2	73.6	100.0
Main Big City	23.1	27.9	49.0	100.0
Urban Fringe	12.7	29.1	58.2	100.0
Medium City	17.5	25.2	57.3	100.0
Small Places	17.2	27.2	55.6	100.0
Age 17				
Nation	12.0	21.8	66.2	100.0
Extreme Rural	11.6	24.6	63.8	100.0
Low Metro	23.9	27.5	48.6	100.0
High Metro	7.0	18.1	74.9	100.0
Main Big City	15.4	22.9	61.7	100.0
Urban Fringe	9.3	21.4	69.3	100.0
Medium City	11.5	20.3	68.2	100.0
Small Places	10.7	21.3	68.0	100.0

subsequent assessments of a learning area as providing both change data--from exercises that were administered in more than one assessment--and baseline data--from exercises that were administered in more than one assessment and exercises that were being administered for the first time in that assessment--which provides a picture of the current status of achievement in that learning area. It is in the sense of "the current status" of achievement at a given point in time that the term "baseline" is being used in this paper.

Within the assessment of any learning area, a large number of exercises are administered to measure student performance. The early reports of assessment data provided detailed information about student performance for individual exercises. This led to complaints that one "couldn't see the forest because of the trees". To remedy this, National Assessment developed data which summarize student performance into a single number for the learning area. This single number is known as either the median or mean National p-value (percentage). It is computed by determining the mean or median for <u>all</u> exercises administered in that learning area at that age level. NAEP also computes the difference between the performance of each of the reporting groups and the nation for a given exercise; this is known as a delta p-value. By determining the mean or median of all the delta p-values for each reporting group, a measure of that reporting group's average difference for the nation is determined. This is presented as the mean or median delta p-value. The data presented in this section of the paper are expressed as either median p-values and median delta p-values or mean p-values and mean delta p-values. While this provides a summary of a great deal of data, it has also led to some rumblings that one "cannot see the trees because of the forest". National Assessment continues to provide detailed exercise by exercise analyses for those desiring it.

Table 6 presents baseline data from the first four years of the assessment for the nation and the STOC categories; Table 7 presents baseline data from the second four years of data collection for the nation and STOC categories. A positive mean or median delta p-value indicates the performance above the national level. A negative mean or median p-value indicates performance below the national level. A quick glance at the tables indicates that high metro students are performing at above the national level in each of the learning areas at all ages while low metro students are performing below the national level in each of the learning areas at all three age levels. When one examines the performance of rural students, one finds that they are generally below that of the nation. But a closer examination also reveals that when one considers their performance in subsequent assessments of the same learning area, they appear to be narrowing the gap between themselves and the national level of performance at each of the age levels. For example, in the first assessment of science, the median delta p-values were -6.3, -6.2, and -3.6 for rural students at ages 9, 13, and 17 respectively (Table 6); in the second assessment of science, the mean delta p-values for rural students were -3.0, -2.2, and -1.6 respectively (Table 7); in the third assessment of science, the mean delta p-values for rural students were 1.4, 0.2, and -0.5 respectively (Table 7). Similar patterns can be traced in writing, reading, citizenship, and social studies.

TABLE 6. Baseline Assessment Data for the Nation and STOC Categories, 1969-1973

	Science	Writing	Citizen-ship	Reading	Literature	Music	Social Studies	Mathe-matics
National Median p-value - Age 9	68.2	28.3	64.1	70.4	43.9	53.8	72.2	36.7
Median delta p-value* - Age 9								
Extreme Rural	-6.3	-4.6	-3.3	-4.4	-3.6	-2.7	-2.8	-3.6
Low Metro	-15.1	-14.2	-5.7	-14.3	-9.4	-8.8	-11.1	-10.8
High Metro	7.2	5.8	3.4	8.4	7.5	5.5	6.7	8.1
Main Big City	-2.7	-2.9	-0.3	1.4	-0.5	0.0	-0.2	-0.9
Urban Fringe	2.6	2.4	0.4	2.1	2.4	0.3	0.6	2.4
Medium City	0.8	2.1	1.4	0.1	-1.2	0.5	0.2	0.8
Small Places	0.9	-0.6	0.0	-0.6	0.1	0.1	0.5	-0.5
National Median p-value - Age 13	58.3	55.4	63.1	68.1	53.4	48.9	66.2	51.3
Median delta p-value* - Age 13								
Extreme Rural	-6.2	-6.3	-4.3	-3.9	-3.1	-1.4	-2.6	-2.1
Low Metro	-13.7	-10.5	-6.0	-8.1	-5.4	-4.5	-8.7	-14.9
High Metro	6.2	7.5	4.3	5.6	5.5	3.4	7.3	10.2
Main Big City	-3.9	-0.4	0.0	-1.3	-0.8	-1.0	-0.2	-1.0
Urban Fringe	2.8	1.8	0.8	2.2	1.3	0.0	0.7	1.5
Medium City	1.9	1.8	1.0	0.4	0.0	0.3	0.5	0.5
Small Places	0.5	-0.7	-0.3	-0.5	-0.6	0.1	-0.5	-0.1
National Median p-value - Age 17	47.0	62.5	61.8	77.5	61.3	49.2	73.8	57.1
Median delta p-value * - Age 17								
Extreme Rural	-3.6	-4.1	-4.8	-2.6	-2.4	-1.3	-2.1	-4.1
Low Metro	-7.4	-10.4	-4.8	-7.7	-7.2	-3.3	-6.3	-14.0
High Metro	5.1	6.6	5.3	5.6	5.3	3.5	6.5	9.9
Main Big City	0.2	-0.6	0.3	1.3	-0.6	-0.4	-0.9	-2.4
Urban Fringe	0.9	3.0	2.4	1.2	1.1	0.5	1.0	1.8
Medium City	1.2	1.6	1.1	0.8	0.2	0.5	0.9	1.8
Small Places	-1.5	0.0	-1.9	-1.4	-0.7	-0.5	-0.4	0.3

* The median delta p-value is a single number used to describe a group's performance. The percent correct (p-value) for an exercise (item) can be expressed for the Nation as a whole or for any of the reporting groups. The difference between the percentage for a group and the Nation is the delta p-value. The median delta p-value for a group is the median of all delta p-values for the group.

NOTE: Data are for the following years 1969-70: Science, Writing, Citizenship
1970-71: Reading, Literature
1971-72: Music, Social Studies
1972-73: Mathematics

TABLE 7. Baseline Assessment Data for the Nation and STOC Categories, 1973-1977

	Science[1]	Career and Occupational Development	Writing	Reading	Art	Citizen-ship	Social Studies	Science
National Mean p-value - Age 9	59.4	65.0	47.3	65.2	40.9	62.1	63.3	50.7
Mean delta p-value - Age 9								
Extreme Rural	-3.0	-3.0	-1.3	-2.2	-1.5	-2.9	-1.8	1.4
Low Metro	-13.4	-10.9	-12.1	-10.0	-5.1	-8.9	-9.8	-11.7
High Metro	7.1	6.6	6.3	6.6	5.0	3.9	4.4	7.6
Main Big City*	-1.5	-1.2	-1.8	-2.3	-0.9	-1.8	-2.0	-2.0
Urban Fringe*	1.6	1.1	0.9	2.1	2.2	2.5	2.6	2.8
Medium City*	1.5	0.6	0.9	0.3	-0.1	0.2	0.3	-1.0
Small Places*	0.7	0.7	0.8	0.6	-0.8	-0.4	-0.2	0.0
National Mean p-value - Age 13	58.3	71.8	51.9	60.7	49.6	63.2	62.9	49.1
Mean delta p-value - Age 13								
Extreme Rural	-2.2	-2.3	-2.9	-3.9	-3.6	-0.7	-0.8	0.2
Low Metro	-13.5	-7.9	-8.7	-11.8	-3.2	-5.5	-6.1	-11.1
High Metro	7.1	5.8	6.3	7.9	3.8	6.6	6.6	6.3
Main Big City*	-2.5	-0.2	1.9	-0.7	-1.4	-1.0	-1.5	-1.2
Urban Fringe*	0.8	1.5	2.4	2.3	2.1	2.9	3.0	0.8
Medium City*	1.2	-0.9	-1.7	-1.2	-0.5	-0.8	-0.8	0.6
Small Places*	1.3	0.2	-0.3	0.7	0.0	-0.4	-0.3	0.1
National Mean p-value - Age 17	42.3	72.8	54.1	72.0	53.0	67.4	67.6	53.5
Mean delta p-value - Age 17								
Extreme Rural	-1.6	0.2	0.6	-1.7	-2.8	-0.1	-0.3	-0.5
Low Metro	-8.2	-6.4	-6.5	-9.2	-2.7	-5.8	-6.1	-12.3
High Metro	4.6	3.4	3.7	6.7	4.6	4.2	4.2	4.4
Main Big City*	-2.8	0.4	2.1	-0.1	0.7	-1.2	-1.2	-2.6
Urban Fringe*	0.0	0.9	0.3	2.2	1.8	0.8	0.8	1.4
Medium City*	0.4	0.6	0.0	0.0	-0.2	0.2	-0.2	1.7
Small Places*	1.2	0.2	0.3	0.6	-0.7	0.2	0.2	0.6

* For Citizenship, Social Studies and Science[2] the data presented are for Big City, Fringes around Big Cities, Medium City, and Smaller Places respectively. These are size-of-community categories rather than size-and-type of community categories.

NOTE: Data are from the following years: 1972-73: Science[1]
1973-74: Career and Occupational Dev., Writing
1974-75: Reading, Art
1975-76: Citizenship, Social Studies
1976-77: Science[2]

4. ASSESSMENT CHANGE DATA

The previous section presents current status or baseline data from the first eight years of data collection. This section will deal only with data from the measurement of changes in performance over time for the areas of science, reading, citizenship, and social studies. To measure change, National Assessment keeps approximately one-half of the exercises from a learning area secure and readministers the exercises as a part of the next assessment of the learning area. Care is taken to replicate the administration of the change exercises as closely as possible. For example, if the exercise contains a typographical mistake or the announcer misread the exercise on the paced audio tape, the "mistake" is replicated. To correct such "mistakes" could effect performance which in turn could result in measuring the effect of correcting the "mistake" rather than true changes in performance on the exercise. To compute change data, National Assessment summarizes the performance on the exercises used to measure change from the prior assessment and compares it to the present data on the same exercises. (This is not meant to imply that NAEP does not compute change data for each exercise; this is computed, used as the basis for computing change measure and is available to the public and researchers.)

The first learning area in which National Assessment measured changes in performance over time was science. The initial assessment of science occurred during the 1969-70 school year; the second assessment of science--and the first measurement of change--was conducted during the 1972-73 school year; and the third assessment of science--and second measurement of change--was administered during the 1976-77 school year.

The change data for science are presented in Table 8. The mean national p-values for each year at each age are presented; beneath the mean national p-value are the mean delta p-values for the type of community and size of community variables. Under the column labeled "change" is the difference between the mean national p-values or the mean delta p-values. An asterisk (*) indicates a mean delta p-value or difference greater than or equal to twice its standard error; these are considered to be statistically significant at the .05 level.

As can be seen from the data, rural students generally performed below the national level for all of the change data. The exception to this is the performance of rural 9-year-olds in 1976-77, where the mean delta p-value is 0.7. What is interesting is the relative improvement in performance by rural students in each measurement of change. The changes in mean delta-p values for the first measurement of change were 1.0, 2.3, 1.6 at ages 9, 13, and 17 respectively; the changes for the second measurement of change were 2.9, 1.6, and 0.5, with the change for 9-year-olds (2.9) being statistically significant. As was indicated by the trends found in the baseline data, rural students are improving in the area of science. A brief examination of science performance by 17-year-olds on the third science assessment materials indicated that rural students tend to perform above the national level on practical, "common-sense" types of science exercises while they perform below the national level on more academically oriented exercises.

TABLE 8. Science Performance, 1969-1977

	1969-70	1972-73	Change	1972-73	1976-77	Change
Age 9						
Nation	61.0	59.8	-1.2*	52.3	52.2	-0.1
Extreme Rural	-3.7*	-2.6*	1.0	-2.2*	0.7	2.9*
Low Metro	-15.2*	-13.4*	1.8	-12.0*	-11.2*	0.8
High Metro	8.1*	-6.6*	-1.5	5.7*	7.3*	1.6
Big City	-3.6*	-3.5*	0.1	-3.6*	-4.6*	-1.0
Fringes-around Big Cities	4.1*	2.8*	-1.3	2.5*	4.2*	1.7
Medium City	1.2*	1.3	0.2	2.5*	-0.7	-3.2*
Smaller Places	-0.1	0.0	0.1	-0.1	0.1	0.2
Age 13						
Nation	60.2	58.5	-1.7*	54.5	53.8	-0.7
Extreme Rural	-4.3*	-2.0	2.3	-1.9	-0.4	1.6
Low Metro	-11.9*	-13.1*	-1.2	-10.7*	-11.6*	-0.8
High Metro	6.4*	6.8*	0.4	5.4	5.6	0.2
Big City	-3.5*	-3.8*	-0.3	-3.1*	-3.2*	-0.1
Fringes-around Big Cities	3.1*	2.0*	-1.1	1.5*	2.5*	1.0
Medium City	0.8	0.3	-0.5	0.1	-0.1	-0.2
Smaller Places	0.0	0.1	0.1	0.6	0.2	-0.4
Age 17						
Nation	45.2	42.5	-2.8*	48.4	46.5	-1.9*
Extreme Rural	-2.9*	-1.4*	1.6	-0.8	-0.3	0.5
Low Metro	-5.1*	-7.3*	-2.3	-8.1*	-10.1*	-2.1
High Metro	5.9*	4.4*	-1.5	4.7*	4.4*	-0.3
Big City	-1.8*	-3.3*	-1.5	-3.6*	-4.4*	-0.8
Fringes-around Big Cities	2.1*	1.3	-0.8	1.1	2.5*	1.4
Medium City	0.7	-0.1	-0.8	-0.1	0.2	0.3
Smaller Places	-0.5	0.5	1.0	0.8*	0.4	-0.4

* Denotes differences or changes in differences greater than or equal to two standard errors.

The first assessment of reading was conducted during the 1970-71 school year; the second assessment of reading--and the first measurement of change--during the 1974-75 school year. The reading assessments included exercises dealing with literal comprehension, inferential comprehension, and references skills. The data from the reading assessments are presented in Table 9. The same conventions as were used with Table 8 are used here.

Overall, the 9-year-olds showed a significant improvement from the first to the second assessment while the performance of 13- and 17-year-olds remained essentially unchanged. For rural students, the results were mixed. At ages 9 and 17, positive changes were found (0.88 and 1.33, respectively) through the results at age 13 show virtually no change.

During the spring of 1974, National Assessment conducted a special survey for the Right-to-Read Effort to examine the functional literacy of 17-year-old students. Right-to-Read consultants reviewed all of the exercises used by NAEP to assess reading and selected a subset of these items which they felt measured practical, everyday types of reading skills that one would need to function in contemporary society. The exercises were administered to a regular NAEP sample of 17-year-olds during the assessment of 17-year-olds that spring. Because many of the exercises had been administered to 17-year-olds as a part of the first assessment of reading, it is possible to look at this data in terms of the measurement of changes in performance over time. The data are presented in Table 10.

As can be seen from the data, the ability of 17-year-olds nationwide and for all of the size-and-type of community categories to respond to functional, everyday types of reading tasks increased between 1971 and 1974. The rural students had the largest change of the STOC categories and were performing at approximately the same level as the nation.

TABLE 9. Reading Performance, 1970-1975

	1970-71	1974-75	Change
Age 9			
Nation	63.98	65.20	1.22*
Extreme Rural	-3.09*	-2.21	0.88
Low Metro	-11.22*	-9.97*	1.25
High Metro	7.59*	5.99*	-1.60
Main Big City	1.12	-2.32*	-3.44*
Urban Fringe	2.51*	2.15*	-0.36
Medium City	-0.02	0.34	0.36
Small Places	-0.54	0.57	1.11
Age 13			
Nation	60.60	60.74	0.14
Extreme Rural	-3.86*	-3.95*	-0.09
Low Metro	-9.59*	-11.84*	-2.25
High Metro	7.71*	7.92*	0.21
Main Big City	1.56	-0.66	-2.22
Urban Fringe	1.99	2.33*	0.34
Medium City	-0.58	-1.17	-0.59
Small Places	-0.60	0.71	1.31
Age 17			
Nation	72.12	72.00	-0.12
Extreme Rural	-3.04	-1.71	1.33
Low Metro	-8.20*	-9.18*	-0.98
High Metro	6.75*	6.71*	-0.04
Main Big City	0.10	-0.06	-0.16
Urban Fringe	1.25	2.20*	0.95
Medium City	0.86	-0.06	-0.92
Small Places	-0.77	0.65	1.42

* Denotes differences or changes in differences greater than or equal to two standard errors.

TABLE 10. Functional Literacy Study for the Right-to-Read Effort 1971-74

	1971	1974	Change
Nation	83.7	85.6	1.9*
Extreme Rural	81.5	85.3	3.9*
Low Metro	76.0	79.4	3.4*
High Metro	88.8	90.3	1.5
Main Big City	84.0	86.0	2.0
Urban Fringe	84.8	87.1	2.3
Medium City	84.4	85.7	1.3
Small Places	83.2	85.5	2.3*

* Denotes changes greater than or equal to two standard errors.

The first assessment of citizenship was conducted during the 1969-70 school year; the second assessment--and the first measurement of change--was conducted during the 1975-76 school year. The data are presented in Table 11.

TABLE 11. Citizenship Performance, 1969-1976

	1969-70	1975-76	Change
Age 9			
Nation	61.29	53.79	-7.50*
Extreme Rural	-5.37*	-2.61	2.76
Low Metro	-14.29*	-8.06*	6.23*
High Metro	5.46*	6.81*	1.35
Big City	-3.91*	-0.69	3.22*
Fringes-around Big City	3.03*	4.49*	1.46
Medium City	1.52	0.18	-1.34
Smaller Places	0.19	-1.45*	-1.64
Age 13			
Nation	64.78	62.15	-2.63*
Extreme Rural	-6.12*	-1.14	4.98*
Low Metro	6.00	9.36*	3.36*
Big City	-2.61*	-3.25*	1.86
Fringes-around Big City	3.19*	4.07*	1.27
Medium City	1.00	-1.77	2.25
Smaller Places	-0.33	-0.07	0.73
Age 17			
Nation	70.03	63.17	-6.86*
Extreme Rural	-7.71*	0.78	8.49*
Low Metro	-5.96*	-9.41*	-3.45
High Metro	4.61*	4.03*	-0.58
Big City	-1.15	-3.55*	-2.40
Fringes-around Big City	2.11*	0.90	-1.21
Medium City	1.32	-0.18	-1.50
Smaller Places	-0.82	0.81	-1.63*

* Denotes differences or changes in differences greater than or equal to two standard errors.

Overall, there were national declines in citizenship performance at all three ages--7.50 at age 9, 2.63 at age 13, and 6.86 at age 17. Yet the rural students narrowed the gap between themselves and the national level at all three ages. Indeed, the changes in the mean delta p-values for rural 13 and 17-year-olds were statistically significant.

The first assessment of social studies was administered during the 1971-72 school year; the second assessment of social studies was conducted with the citizenship assessment during the 1975-76 school year. The data are presented in Table 12.

At the national level, there was a slight increase at age 9, a slight decline at age 13, and a significant decline at age 17. Among rural students, the 9-year-olds essentially closed the gap between themselves and the nation while 13 and 17-year-olds narrowed the gap between themselves and the nation. A closer examination of the performance of rural 9-year-olds revealed that there was a slight increase in their performance on the social studies knowledge exercises, a significant increase in their performance on social studies skills exercises (which dealth with obtaining and interpreting information) and a decrease in their performance on the social studies attitudes exercises.

TABLE 12. Social Studies Performance, 1971-1976

Age 9	1971-72	1975-76	Change
Nation	65.79	66.98	1.19
Extreme Rural	-1.20	-0.07	1.13
Low Metro	-12.61*	-10.75*	1.86
High Metro	7.35*	6.01*	-1.34
Big City	-4.87*	-1.95*	2.92*
Fringes-around Big Cities	2.86*	3.03*	0.17
Medium City	0.54	0.17	-0.37
Smaller Places	0.15	-0.14	-0.29
Age 13			
Nation	60.52	59.35	-1.17
Extreme Rural	-3.04*	-1.43	1.61
Low Metro	-8.57*	-8.23*	0.34
High Metro	6.77*	7.37*	0.60
Big City	-1.03	-2.71	-1.68
Fringes-around Big Cities	2.24*	3.84*	1.60
Medium City	0.72	-0.90	-1.62
Smaller Places	-1.30*	-0.25	1.05
Age 17			
Nation	71.32	68.14	-3.18*
Extreme Rural	-2.63*	-1.40	1.23
Low Metro	-7.66*	-7.50*	0.16
High Metro	6.86*	4.76*	-2.10*
Big City	-3.03*	-1.66	1.37
Fringes-around Big Cities	3.27*	1.68	-1.59
Medium City	0.82	-0.26	-1.08
Smaller Places	-0.97*	0.03	1.00

* Denotes differences or changes in differences greater than or equal to two standard errors.

8. SUMMARY AND RECOMMENDATIONS

The purpose of this paper has been to examine the performance of rural students in terms of National Assessment data--both baseline and change data. The data presented in the previous two sections reveal that although rural students have traditionally performed below the national level, the trend from the baseline data is that rural students are improving and in some instances, have reached the national performance level; the change data substantiates this trend. However, it must be remembered that given the time and space limitations surrounding this paper, it has only been possible to look at the summary data (or, "the forests"). By examining the exercise by exercise level data, it would be possible to isolate the strengths and weaknesses of rural students in the various learning areas (or, to find the "good" and the "blighted" trees). National Assessment recognizes this limitation, and in response to the limitation, it makes available its data base for secondary analyses. Thus, it would recommend that secondary analyses be funded to examine the data base for the strengths and weaknesses of rural students.

As was noted in Section 2, National Assessment neither identifies nor institutes special procedures to guarantee the inclusion of the children of migrant workers. This is a definite limitation of the data base. This writer knows of no national data dealing specifically with the performance of migrant children. This leads to the next recommendation. National Assessment has in the past conducted special purposes studies as an "add-on" to its regular data collection efforts. The assessment of functional literacy conducted for the Right-to-Read Effort is one example of this; the special assessment of the performance and participation of women in mathematics currently being conducted

for the National Institute of Education is another example. Thus, it is recommended that the federal government explore the possibility of conducting an assessment of migrant children.

The final recommendation deals with the limitations of the existing National Assessment data base on rural students. At present, rural students account for approximately 10 percent of the sample at each age level. While this is sufficient for reporting the performance of rural students, it is not adequate to examine the performance of rural students by region, of rural students by race, of rural students by socioeconomic variables. Thus, it is recommended that National Assessment be provided with the resources necessary to increase the samples size for rural students to provide more detailed information about their educational achievement.